Melody contour identification and instrument recognition using semitone mapping in Nucleus Cochlear Implant recipients

Omran, S A ; Lai, W K ; Dillier, N

Abstract: Cochlear implants (CIs) were originally aimed at restoring speech perception for patients with profound hearing loss. Many postlingually deafened CI patients report that music is not well perceived while others enjoy it. Music consists of complex sounds composed of tones with harmonic structure of overtones and temporal fine structure. The harmonic structure is not preserved by the current standard (Std) ACE (advanced combination encoders) mapping and the temporal fine structure is not well encoded. The mapping is believed to produce distortion due to compression oin the low frequency range. In 2008 we proposed two new semitone (Smt) mappings (Smt-LF and Smt-MF) in two frequency ranges (130-1502 Hz and 440-5040 Hz) respectively (Omran et al. 2008). Smt mapping is expected to preserve the harmonic structure representation of overtones and this may improve melody recognition with CIs. In this paper two psychoacoustic experiments (melody contour identification (MCI) (Galvin et al. 2007) and instrument recognition (IR)) were conducted with three different conditions (Std, Smt-MF and Smt-LF mappings) with CI recipients by streaming processed stimuli directly to the implant. The MCI test included five patterns (rising - rising falling - flat - falling rising – falling). Each pattern consisted of five tones; each tone had a fundamental frequency and four overtones. The lowest fundamental frequency of each pattern is called "root". Signals had two different roots A3 (220 Hz) and A4 (440 Hz). Proposed nine patterns with three roots (A3, A4 and A5) by Galvin et al. (2007) were examined in a pilot test. This test took a long time and the preliminary results showed a possibility to reduce the number of patterns to five and eliminate the fifth octave root (A5). In the IR test, four pairs of instruments (Trumpet and Trombone, Flute and Clarinet, Violin and Cello, Guitar and Piano) from four groups (Brass, Woodwind, Struck and String instruments) respectively were used. MCI and IR tests were conducted with 8 CI recipients. Results from MCI tests showed an improvement with Smt mapping in respect to Std mapping or at least similar results. However, wrong identification occurred with notes having filtered out partials. CI recipients showed an increase in identifying melody contour patterns with Smt mappings. Instrument identification performance decreased with semitone mappings.

Originally published at:
Omran, S A; Lai, W K; Dillier, N (2009). Melody contour identification and instrument recognition using semitone mapping in Nucleus Cochlear Implant recipients. In: 12. Jahrestagung der Deutschen Gesellschaft für Audiologie, Innsbruck, 11 March 2009 - 14 March 2009. DGA e.V., 1-6.

Melody contour identification and instrument recognition using semitone mapping in Nucleus Cochlear Implant recipients

Sherif A. Omran [1,2], Waikong Lai [1], Norbert Dillier [1]

[1] Laboratory of Experimental Audiology, University Hospital Zurich - Switzerland
[2] Institute of Neuroinformatics – University of Zurich- Switzerland

Abstract

Cochlear implants (CIs) were originally aimed at restoring speech perception for patients with profound hearing loss. Many postlingually deafened CI patients report that music is not well perceived while others enjoy it. Music consists of complex sounds composed of tones with harmonic structure of overtones and temporal fine structure. The harmonic structure is not preserved by the current standard (Std) ACE (advanced combination encoders) mapping and the temporal fine structure is not well encoded. The mapping is believed to produce distortion due to compression oin the low frequency range.

In 2008 we proposed two new semitone (Smt) mappings (Smt-LF and Smt-MF) in two frequency ranges (130-1502 Hz and 440-5040 Hz) respectively (Omran et al. 2008). Smt mapping is expected to preserve the harmonic structure representation of overtones and this may improve melody recognition with CIs. In this paper two psychoacoustic experiments (melody contour identification (MCI) (Galvin et al. 2007) and instrument recognition (IR)) were conducted with three different conditions (Std, Smt-MF and Smt-LF mappings) with CI recipients by streaming processed stimuli directly to the implant.

The MCI test included five patterns (rising - rising falling - flat - falling rising – falling). Each pattern consisted of five tones; each tone had a fundamental frequency and four overtones. The lowest fundamental frequency of each pattern is called “root”. Signals had two different roots A3 (220 Hz) and A4 (440 Hz). Proposed nine patterns with three roots (A3, A4 and A5) by Galvin et al. (2007) were examined in a pilot test. This test took a long time and the preliminary results showed a possibility to reduce the number of patterns to five and eliminate the fifth octave root (A5). In the IR test, four pairs of instruments (Trumpet and Trombone, Flute and Clarinet, Violin and Cello, Guitar and Piano) from four groups (Brass, Woodwind, Struck and String instruments) respectively were used.

MCI and IR tests were conducted with 8 CI recipients. Results from MCI tests showed an improvement with Smt mapping in respect to Std mapping or at least similar results. However, wrong identification occurred with notes having filtered out partials. CI recipients showed an increase in identifying melody contour patterns with Smt mappings. Instrument identification performance decreased with semitone mappings.

Introduction

Cochlear Implants (CIs) were originally aimed at restoring speech perception for patients with profound hearing loss. Many post lingual CI patients report that music is not well perceived. Music consists of complex sounds composed of tones with harmonic structure of overtones. Most musical instruments generate fundamental frequencies below 1kHz (Pierce 1983). We proposed two semitone (Smt) mappings to improve melody representation with CI patients (Omran et al. 2008). Smt-LF and Smt-MF mappings cover frequency ranges from 130 to 1502 Hz and from 440 to 5040 Hz respectively. Smt mappings preserve the representation of the harmonic structure of musical tones. This may improve music perception.

Music has different aspects such as pitch, melody and timbre. Psychoacoustic tests evaluate these various dimensions. Frequency representation, loudness and temporal resolution are main factors that affect perception and understanding (Laneau and Wouters 2004). To examine music perception with semitone (Smt) mapping in this study; two psychoacoustic tests (melody contour identification (MCI) (Galvin et al. 2007) and instrument recognition (IR)) were conducted in three testing conditions (Standard (Std) ACE (advanced combination encoders), Smt-LF and Smt-MF mappings) with CI recipients.

Melody is one of the important aspects of music (Sadie and Grove 1995). It can be described as a group of tones perceived as a single entity (Terhardt 1998). Each tone has a harmonic structure of overtones and preserving this structure may improve melody perception. Timbre (tone color) is another aspect of music. Different instruments are characterized by their timbre (Helmholtz 1954). Timbre depends on the relationship between intensities of different partials. In the IR test, patients heard sounds from different musical instruments using different mappings and had to identify the instrument by which the sounds were played.

This paper is organized as follows: Section (1) describes the hypothesis of the two experiments (MCI and IR). Section (2) illustrates methods and procedures of each test. Section (3) shows the results of each test. Section (4) is a general discussion for the three experiments. The conclusion is located in section (5).

1. Hypothesis

- Smt mapping may provide higher MCI scores than Std mapping due to preserving the harmonic structure. Ambiguities may occur with Smt-MF mapping at low frequencies due to filtering out partials below 440 Hz and performance may decrease with Smt-LF mapping because frequencies are transposed to higher ranges.
- Improving frequency representation with Smt mapping might improve instrument recognition compared to the Std mapping.

2. Methods and procedures

2.1 Experiment 2: Melody Contour Identification (MCI)

A root note (A) from three (3rd, 4th and 5th) octaves was utilized to build up five patterns (rising – rising falling – flat – falling rising –falling) (Galvin et al. 2007). Each pattern was composed of five complex synthetic tones and each tone had five harmonic partials. The amplitude of each partial decreased 20% from the previous one. To avoid loudness and temporal cues all tones were designed to have similar temporal structure and the RMS of each pattern was normalized at -15 dB. Each pattern had a duration of 250 ms with a 50 msec pause in between tones. Tones were faded out/in 10 ms at the beginning and the end respectively. The fundamental frequency of each pattern is the same as the root note. Galvin et al. (2007) designed nine patterns. However, initial testing showed no significant difference between some patterns and therefore only five patterns were utilized to test CI subjects as shown in Figure (1).

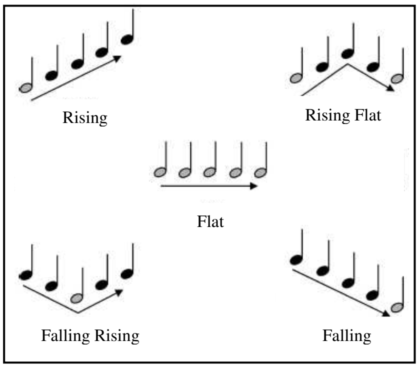

Figure (1): Showing different patterns used in the MCI test with CI recipients. Root notes are filled with gray.

The semitone interval between successive tones was varied between 1 and 3 semitones maximum for each pattern used with octave 3 and 4. This is summarized in Table (1).

CI Recipients			
Interval	1	2	3
Octave A4	☑	☑	☑
Octave A3	☑	☑	☑

Table (1): Showing semitone intervals between successive notes used with NH subjects (left) and CI recipients (Right).

Different patterns were processed by a Matlab application called "*Checker*". Patient's maps were first loaded from a clinical database. The wave files for the different patterns were loaded and processed. The Checker processed files using 22 channels and generated electric stimuli for each file using the Nucleus Matlab Toolbox (NMT). Additionally, it analyzed stimulus files for different sorts of errors. Electrodograms of stimuli presented acoustically via loudspeaker and electrically via direct electric input to the sound processor were captured and calibrated. The stimuli were set to 65dBA equivalence. Once stimulus files were approved, they were converted into Xml files and stored. Xml files were streamed to the Freedom CI sound processor from the Macarena test environment using the Nucleus implant communicator (NIC) software from Cochlear Corporation.

2.2 Experiment 3: Instrument recognition

Recorded music pieces played with eight instruments by professional musicians were used as test material. The four instrument groups (Brass, Woodwind, Struck and String) with two instruments per group are illustrated in Figure (2).

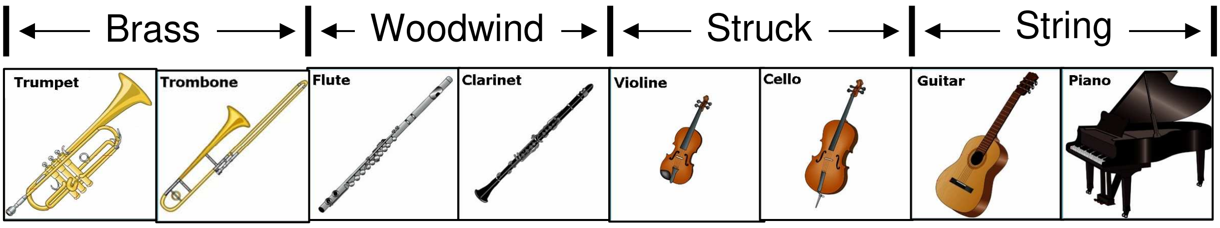

Figure (2): Two instruments from four instrument groups (Brass, Woodwind, Struck and String) used in the instrument recognition test.

The average root mean square (RMS) of sound intensity from different instruments was normalized to -15 dB level. The signals were processed with three (Std – Smt-MF and Smt-LF) mappings using patient specific settings of minimum and maximul current levels per electrode retrieved from a clinical database. The Checker was used to process the different sounds and prepare XML files for each patient. Patients were seated in front of a touch screen and an XML file was streamed to the speech processor from Macarena test environment using NIC. Patients had to select an instrument corresponding to the perceived sound. 8 adult post-lingual patients performed the test. All subjects had at least 1 year experience using a CI device.

Before testing, patients were trained on different sounds in familiarization and feedback sessions. In a familiarization session, the patient selects a button on the screen and hears the corresponding sound. In a training session with feedback, false selections are indicated on the screen and the same sounds can be repeatedly presented.

3. Results

3.1 Melody contour identification

In the MCI test different contour patterns are presented to CI recipients. The mean correct identification scores of the MCI test is calculated for different octaves and different semitone intervals using Std, Smt-MF and Smt-LF mappings.

8 CI recipients conducted the MCI test with the same number of repetitions and the same mapping conditions.

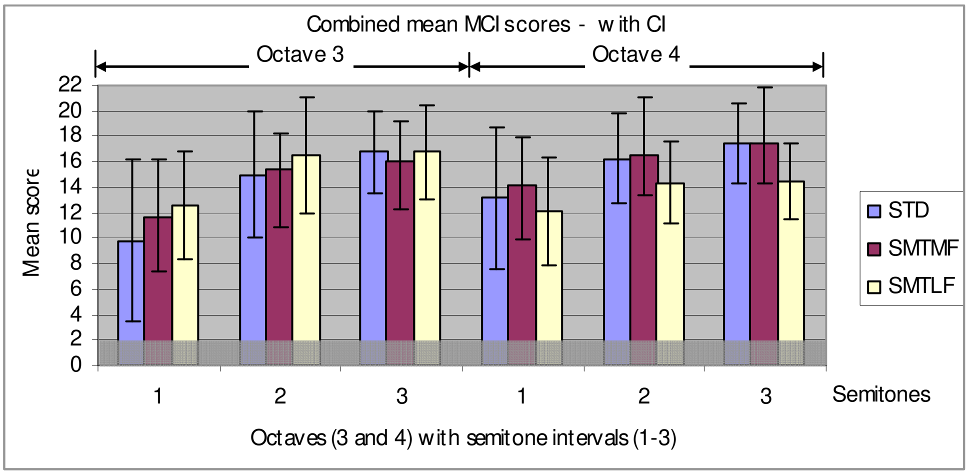

Figure (3): Results with standard (blue), semitone (Smt-MF) (dark red) and (Smt-LF) (yellow) mappings. Two octaves (3 and 4) are plotted with semitone intervals from 1 to 3. Chance level is shaded in gray.

Figure (3) shows results with Std, Smt-MF and Smt-LF mappings. There is an increase in correct contour identifications with Smt-MF at octave 3 and 4 with semitone intervals 1 and 2. Scores with Std and Smt-MF mappings in interval 3 are almost similar. Results with Smt-LF mapping show that there is a slight increase in identification with Smt-LF at octave 3 with semitone intervals 1 and 2. However, the Smt-LF score decreases at octave 4 compared with Std and Smt-MF mappings. This decrease may be due to filtering out high frequency partials with Smt-LF.

The CI recipient's difficulty to resolve melody contours is shown in Figure (4). A significant decrease in errors identifying melody contours with Smt-MF at octave 3 with 1 interval is was found in comparison with STD mapping. The error was even smaller with Smt-LF mapping.

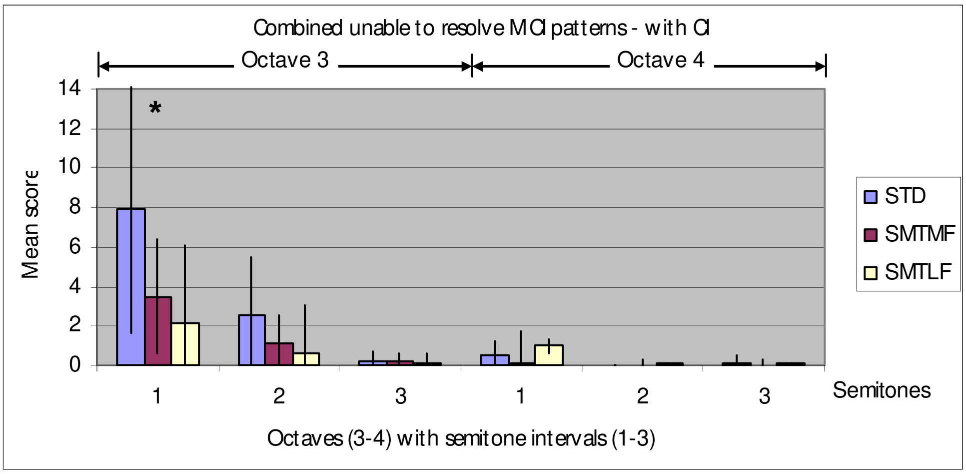

Figure (4): Combined results using standard, Smt-MF and Smt-LF mappings with octaves 3 and 4 using intervals from 1 to 3.

3.2 Instrument recognition

8 CI subjects underwent the IR test. Their task was to recognize an instrument used to play a musical piece. There were eight instruments from four instrument families. The results are analyzed to show the score of correct instrument and family recognition.

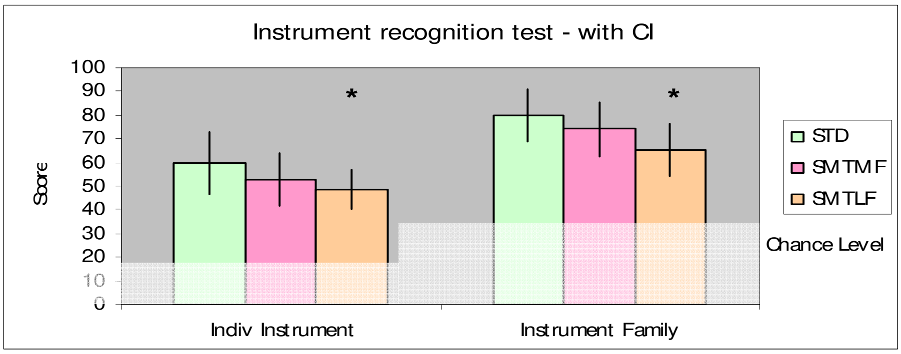

Figure (5): Showing results of individual musical instruments and family recognition test with CI recipients using standard (Std) (blue) and Smt-MF (dark red) and Smt-LF (yellow) mappings. White shadow shows chance level.

Figure (5) describes results from CI recipients to identify an instrument and instrument family with Std, semitone (Smt-MF and Smt-LF) mappings. It illustrates that the identification score decreased significantly with Smt-LF mappings compared to Std mappings. This may be because the timbre is encoded partially by temporal fine structure (Loizou et al. 2003). Since these strategies do not encode the temporal fine structure, patients may be relying on the spectrum to identify different instruments. Since the Std mapping is covering the widest frequency range (180-7800 Hz) compared to semitone mapping Smt-LF and Smt-MF ranges (130-1502 Hz) and (440-5040.5 Hz) respectively; the highest amount of spectral information is transmitted with Std mapping.

Discussion

Smt-MF mapping ameliorated correct pitch identification. However, pitch reversals may have occurred due to filtering out the fundamental frequency which had a negative effect on the identification scores. Smt-MF filters out partials below 440 Hz and above 5040.5 Hz, and Smt-LF mapping filters out partials below 130 Hz and above 1502 Hz. Smt-LF mapping transposes the frequencies into higher bands. This may simplify identifying tones at low frequency octaves (e.g. octave 3) but may affect characterization of perceived tones at high frequencies due to filtering out overtones greater than 1502 Hz. The results nevertheless demonstrated that semitone mapping may improve pitch ranking due to improving frequency representation.

Melody can be described as a group of tones that are perceived as a single entity. Melody contour identification (MCI) may be regarded as pitch ranking of multiple tones. MCI (Galvin et al. 2007) was tested with cochlear implant recipients by presenting different contour patterns. The results of this test showed that there was a significant improvement using Smt-MF and Smt-LF compared to Std mapping at different octaves and intervals. However, with Smt-LF there was a significant decrease in contour identification at octave 5 with normal hearing subjects, this may be because of filtering out high frequency partials with Smt-LF. Some pattern were identified as flat when they were not flat. The results with these particular patterns were further analyzed. The analysis showed that there was a significant decrease in identification error with Smt-MF and Smt-LF mappings. With Smt-LF mapping the error decreased even more than Smt-MF.

Pitch and melody are two aspects of music. Timbre is another aspect that characterizes different instruments (Helmholtz 1954). Timbre depends on the frequency spectrum as well as the temporal fine structure of the perceived sounds. Results with cochlear implant recipients in the instrument recognition test showed that there was a significant decrease in instrument and family recognition with semitone mappings. The decrease is significant with Smt-LF mapping. Semitone mapping is built upon the ACE strategy. This strategy does not encode fine temporal resolution. Thus, patients may have been strongly relying on power spectral density of signals as suggested by (Drennan and Rubinstein 2008). The standard mapping covers a range from 188-7980 Hz, while Smt-LF and Smt-MF cover the frequency range from 130-1502 Hz and 440-5040 Hz respectively. Since the standard mapping has a wider frequency range than the semitone mappings, the root mean square of the encoded spectrum will be higher than with semitone mappings. Another factor may be patient familiarity with semitone mapping compared to the daily used standard mapping.

Conclusion

Melody contour identification (Galvin et al. 2007) was increased with semitone mapping compared with standard mapping. This increase differs between Smt-MF and Smt-LF. Although, Smt-LF mapping provides better pitch ranking and melody identification results, the perceived sounds are highly pitched and unpleasant for some patients. Instrument identification scores were decreased with semitone mapping.

This illustrates a need of further development to encode the temporal fine structure for improved perception of music features.

Acknowledgement

This project was supported by Swiss national foundation grant number 320000-110043. Thanks to Michael Büchler for his support during the earlier stages of this project.

References

Drennan WR, Rubinstein JT (2008) Music perception in cochlear implant users and its relationship with psychophysical capabilities. J Rehabil Res Dev 45: 779-789

Galvin JJI, Fu Q-J, Nogaki G (2007) Melodic Contour Identification by Cochlear Implant Listeners. Ear and Hearing 28: 302-319

Helmholtz HLF (1954) On the Sensations of Tone as a Physiological Basis for the Theory of Music. Dover, New York

Laneau J, Wouters J (2004) Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear impantees. Acoustical Society of America 116: 3606-3619

Loizou PC, Stickney G, Mishra L, Assmann P (2003) Comparison of speech processing strategies used in the Clarion implant processor. Ear Hear 24: 12-19

Omran S, Büchler M, Lai W, Dillier N (2008) Pitch perception using a Semitone mapping to improve music representation in Nucleus Cochlear Implants. In: DGA Congress, Kiel, Germany

Pierce J (1983) The science of musical sound. Scientific American Books

Sadie S, Grove G (1995) The New Grove Dictionary of Music and Musicians. In, vol 17. Grove, London

Terhardt E (1998) Akustische Kommunikation. Springer-Verlag, Berlin